THE EFFECT OF PROVIDING CLEAR AIR TURBULENCE ASSESSMENTS TO COMMERCIAL AIRLINE PILOTS

Ellen J. Bass and Diego J. Castaño
Search Technology, Inc.
Norcross, GA

ABSTRACT

The current turbulence assessment and reporting system is subjective and depends upon verbal communications on the radio. We prototyped a system that calculates and displays objective turbulence metrics. In a full motion simulator experiment, commercial airline pilots using objective turbulence metrics were able to provide significantly better passenger and flight attendant safety, passenger comfort, and fuel economy than those using the current convention. Pilots with displays of surrounding traffic's turbulence spent less time on the radio and made significantly better decisions with respect to passenger safety, passenger comfort, and fuel economy than those relying on radio transmissions.

INTRODUCTION

Airborne encounters with clear air turbulence (CAT) affect the safety and efficiency of commercial aviation (FAA, 1994). Currently the best available real-time information concerning CAT comes from pilot reports (PIREPs) verbally communicated over the radio (Bass & Ernst-Fortin, 1999). A shortcoming of the current turbulence reporting system is that few PIREPs are made (Schwartz, 1996). Also pilots base the reports on aircraft response and not on the atmospheric conditions themselves. Also pilot turbulence assessments vary between pilots (Bass, Kvam, & Campbell, in review).

Solving this problem begins with developing an objective way to measure turbulence. The National Center for Atmospheric Research (NCAR) has an approach to map aircraft response to an objective, aircraft independent turbulence metric (Cornman, Morse, & Cunning, 1995). We developed a prototype system that calculates an objective turbulence metric in real-time based on NCAR's approach (Bass, Castaño, et al., 2000).

The notion behind the prototype is to calculate objective turbulence metrics in real-time and to transmit assessments to surrounding aircraft. If traffic turbulence data could be displayed on the flight deck, pilot awareness of atmospheric turbulence should be improved and the necessity for turbulence related communication over the radio lessened. To investigate these hypothesized effects, our prototype included an ownship display of objective turbulence information and a traffic turbulence display (TTD) with metrics from surrounding aircraft. We evaluated the prototype at a commercial airline's training facility. This paper describes the prototype evaluation with a stricter set of rules for the decision quality measures than previous analyses (Bass, Jones, & Castaño, 2000).

EXPERIMENTAL METHOD

Apparatus

The experiment was conducted in a level D certified full motion B767-300ER simulator. Based on verbal PIREP content (Spence, 2001), a pilot survey (Bass & Ernst-Fortin, 1999), a usability study (Castaño & Bass, 2000), and concern over display clutter, current and peak turbulence data were displayed. Although turbulence intensity definitions currently include five levels (i.e., none/smooth, light, moderate, severe, and extreme), pilots create categories in between (e.g., "very light chop") (Bass, et al., in review). Thus a ten-point scale was used for intensity. A three-point scale was used for frequency. The format for turbulence information was d/Cd (where d is a digit (0-9) and C is a character (O, I, C)) (Table 1).

A modification was made to the Engine Indication and Crew Alerting System (EICAS) to display ownship turbulence information. A TTD, running on a laptop placed on the center pedestal between the pilots, displayed traffic turbulence data. Its human-computer interface (HCI) was modeled after Traffic Alert and Collision Avoidance System (TCAS) (Curran, 1992).

Table 1. Turbulence display format

d	/	C	d
Current intensity level of turbulence (0-9)		Peak frequency (O, I, C, or blank) over the past 5 minutes	Peak intensity level of turbulence (0-9) over the past 5 minutes
		O = occasional (< 1/3 of the time)	
		I = intermittent (1/3 –2/3 of the time)	
		C = continuous (> 2/3 of the time)	
		Blank if the turbulence has just started	

The TTD was oriented track up with ownship position at the bottom. The relative altitude of traffic was displayed with two digits (hundreds of feet) plus a vertical speed arrow. Altitude location depended on relative position. Altitude was displayed above and turbulence below the traffic symbols for aircraft at or above ownship altitude (and vice versa for aircraft below ownship). Data for traffic four flight levels (FLs) above and below ownship appeared on the TTD.

Independent measures

Technology intervention treatments. Some independent measures were based on possible technology interventions that could serve as sources of turbulence information:

1. Baseline: The baseline was today's system. Pilots made turbulence reports on the radio using the current negligible/light/moderate/severe/extreme (N/L/M/S/E) convention.

2. Ownship: Ownship meant that only the subject aircraft had the ownship turbulence display on EICAS. Turbulence information must still be verbally communicated as with the baseline.

3. Metric in reports (MIR): MIR envisioned a situation where all aircraft have the ownship display so that all turbulence-related communications include objective metrics.

4. TTD40: TTD meant that the subject aircraft had the ownship display plus a TTD with assessments from aircraft in the area. One version had a maximum range of 40 nautical miles (NM) (similar to current TCAS capabilities). All turbulence-related communications on the radio include objective metrics.

5. TTD120: A second TTD version had a 120 NM range based on expected Automatic Dependent Surveillance Broadcast (ADS-B) capabilities above FL 240 (RTCA, 1998). All turbulence-related communications include objective metrics.

The intervention effects can be investigated. Also, their inherent assumptions can be used to group them to create other dependent measures:

- Traffic turbulence assessment (traffic turbulence information based on a subjective scale vs. automated assessments with an objective scale),

- Traffic turbulence presentation (traffic turbulence report provided verbally vs. display-based),

- Objective traffic turbulence presentation (objective traffic turbulence report provided verbally vs. display-based)

Scenarios. Another independent measure was scenario: a combination of atmospheric condition and traffic. The four atmospheric conditions were:

1. All light - Flying at any altitude would be experienced as light.

2. Prepare cabin - Ownship would first experience light turbulence. Ahead every flight level will be experienced as moderate (intensity level 5) 80-160 NM in front of the aircraft.

3. Smooth above - Ownship would initially experience light turbulence. The turbulence increases to moderate at the ownship's FL directly ahead. The turbulence is negligible above.

4. Smooth below - Ownship would initially experience light turbulence. The turbulence increases to moderate at the current FL directly ahead. The turbulence is negligible below.

There were two traffic mixes with each atmospheric condition. The traffic mixes differed in how they assessed and therefore reported turbulence under the N/L/M/S/E convention. In one traffic mix, traffic reported turbulence in a manner matching what the ownship would experience. In the other, some traffic over- or under-report turbulence intensity.

The eight scenarios were as follows:

- All Light (AL): traffic report light turbulence

- All Light/Over-report (ALO): aircraft directly ahead at ownship's FL and aircraft 4000 feet above ownship report moderate instead of light
- Prepare Cabin (PC): traffic 80-160 NM away report moderate at all levels
- Prepare Cabin/Under-report (PCU): aircraft directly ahead at ownship's FL report light instead of moderate
- Smooth Above (SA): traffic 80-160 NM away report moderate at current FL and smooth above
- Smooth Above/Under-report (SAU): aircraft directly ahead at ownship's FL report light instead of moderate
- Smooth Below (SB): traffic 80-160 NM away report moderate at current FL and smooth below
- Smooth Below/Under-report (SBU):aircraft directly ahead at ownship's FL report light instead of moderate

Dependent measures

Length of radio communications. One measure was ride report radio communications time. The TTD interventions provided traffic turbulence information without pilots having to request ride reports on the radio frequency. However, based on our previous survey results (Bass & Ernst-Fortin, 1999), we were not certain if the TTD40 intervention had sufficient range to have any impact.

Decision-making quality. Decision-making quality was analyzed with respect to passenger and FA safety, passenger comfort, and fuel economy. The measures depended on the scenarios. Using the FAA's suggested turbulence procedures (Table 2), echoed in the participating airline's manuals, we created rules to assess passenger and FA safety decision quality (Tables 3 and 4). We also created rules to assess passenger comfort and fuel economy decision quality (Tables 5 and 6). Tables 3-6 also describe our hypotheses when pilots would make poor decisions (over- or under-reports should mislead the pilots).

Table 2. Turbulence Procedures (FAA, 1994)

Intensity	Response
Light	Seat Belt (S/B) Sign On. Captain notifies lead flight attendant (FA) of intensity and duration. FA makes standard PA; FAs perform visual check of S/Bs.
Mod.	S/B Sign On. Captain makes PA requesting passengers return to seats and ensure S/Bs are fastened. Captain notifies FA of intensity and duration; ensure that cockpit PA heard in cabin.
Severe or greater	S/B Sign On. Captain makes PA instructing passengers and FAs to be seated, followed by interphone call to lead FA. FAs sit down immediately and secure nearest s/b.

Test design

Sixteen pilots were randomly assigned to four groups (ownship, MIR, TTD40, and TTD120), four per group. Each flew ten trials. So that the pilots would realize there were no "tricks", the first two trials for all participants were baseline. The latter eight trials were non-baseline technology ones: one per scenario. For each of the five interventions, there were 32 trials (4 trials of each scenario). Scenario was a within subjects variable while interventions were between subjects (with subjects nested within non-baseline treatments).

Procedure

An experimental session consisted of a briefing, flight experience questionnaire, two baseline trials, technology intervention briefing and training, eight non-baseline trials, and a de-briefing. The longest session was four hours in length. Before any trials, the participant was given flight plan paperwork including the final weather briefing. Each trial was initialized at the same location in cruise. The participant flew the two baseline trials. After a break, the participant received training for the ownship display. If the participant was in a TTD group, he also received that training. After all trials, a debrief occurred.

In all trials, the participant acted as the pilot flying (PF). Each participant was told that the PF is the decision-maker for flight planning purposes and for initialization of communications regarding flight conditions and cabin service considerations. Three confederates participated in the trials. One was the other crewmember, who made all radio communications with air traffic control (ATC) under the participant's direction. A second acted as ATC and as the traffic. If the participant wanted ride reports, ATC first asked traffic at the aircraft's FL. The confederate responded as the traffic would and then, as ATC, repeated the intensity along with the location of the traffic. In baseline and ownship trials, reports were in terms of the N/L/M/S/E convention. In the others, reports were objective. If the participant asked for more reports but did not specify altitude, the confederate solicited and provided information from aircraft at higher FLs. A third query included any other aircraft in the scenario. A third confederate acted as lead FA and as recorder.

Participants

Eight male captains and eight male first officers (FOs) with B767-300ER experience participated. All worked for the same commercial carrier. The captains averaged 13,600 flight hours (ranging 7,500-35,000 hours); FOs averaged 7900 hours (ranging 3700-12,000 hours). All participants were paid volunteers.

RESULTS

Length of Radio Communications

The TTD120 reduced radio communication time. Tests of between-factors effect showed a significant main effect for treatment condition with $F_{(4,15)} = 5.982$, $p = 0.004$. Degrees of freedom were corrected by using the Huynh-Feldt epsilon adjustment for sphericity. Post-hoc multiple comparisons using the Bonferroni adjustment showed significant differences between the TTD120 and all others.

Decision quality

One–sided tests of proportion were carried out for each decision/grouped technology intervention pair with respect to passenger and FA safety, ride quality and economy. The null hypothesis for each test was that the proportion of good decisions for the first group of technology interventions listed in column headers of Table 7 is greater than the second (significant at $p<0.05$ in bold). The Fischer-Irwin test, a conservative method using the hypergeometric distribution, was employed.

Table 3. Passenger Safety Decision Quality Rules

	Decision Quality Rule	Hypothesis in Over/Under-report Scenarios
AL	S/B sign on	Not applicable
ALO		Baseline and ownship: hearing mod. turb. ahead, make PA or descend to smooth air
PC	S/B sign on AND PA made before level 5 turb.	Not applicable
PCU		Baseline and ownship: hearing light turb. ahead, only illuminate S/B sign
SA	S/B sign on AND PA made before level 5 turb. OR S/B sign on AND climb before level 5 turb.	Not applicable
SAU		Baseline and ownship: hearing light turb. ahead, only illuminate S/B sign
SB	S/B sign on AND PA made before level 5 turb. OR S/B sign on AND descend before level 5 turb.	Not applicable
SBU		Baseline and ownship: hearing light turb. ahead, only illuminate S/B sign

Table 4. Flight Attendant Safety Decision Quality Rules

	Decision Quality Rule	Hypothesis in Over/Under-report Scenarios
AL	Not applicable – FA safety is not compromised	Not applicable
ALO		
PC	Make FA call OR make PA with FA-related content.	Not applicable
PCU		Baseline and ownship: hearing light turb. ahead, do not inform FAs
SA	Make FA Call OR Make PA with FA-related content OR Climb before level 5 turb.	Not applicable
SAU		Baseline and ownship: pilots who do not change FL, hearing light turb. ahead, do not inform FAs
SB	Make FA Call OR Make PA with FA-related content OR Descend before level 5 turb.	Not applicable
SBU		Baseline and ownship: pilots who do not change FL, hearing light turb. ahead, do not inform FAs

Table 5. Passenger Comfort Decision Quality Rules

	Decision Quality Rule	Hypothesis in Over/Under-report Scenarios
AL	Not applicable – ride quality is the same regardless of altitude	Not applicable
ALO		
PC		
PCU		
SA	Pilot should request higher	Not applicable

SAU	altitude	Baseline and ownship: hearing light turb. ahead, do not change FL
SB	Pilot should request lower altitude	Not applicable
SBU		Baseline and ownship: hearing light turb. ahead, do not change FL

Table 6. Fuel Economy Decision Quality Rules

	Decision Quality Rule	Hypothesis in Over/Under-report Scenarios
AL	Pilot should stay at current altitude	Not applicable
ALO		Baseline and ownship: hearing mod. turb. ahead and above, descend
PC	Pilot should stay at current altitude	Not applicable
PCU		Baseline and ownship: hearing light turb. ahead, stay at the current altitude
SA	Pilot should request higher altitude	Not applicable
SAU		Baseline and ownship: hearing light turb. ahead and above, do not change FL
SB	Pilot should request lower altitude	Not applicable
SBU		Baseline and ownship: hearing light turb. ahead and above, do not change FL

Table 7. Proportion tests for quality of decisions

	Traffic turbulence assessment: TTD40+TTD120+MIR(objective) vs. Baseline+Ownship(subjective)	Traffic turbulence presentation: TTD40+TTD120 (display) vs. Baseline+Ownship+MIR (verbal)	Objective traffic turbulence presentation: TTD40+TTD120 (display) vs. MIR (verbal)
Passenger Safety	**63/80 (79%) vs. 37/60 (62%) p = 0.0217**	**40/48 (83%) vs. 60/92 (65%) p = 0.0180**	40/48 (83%) vs. 23/32 (72%) p = 0.170
FA Safety	**65/66 (99%) vs. 43/48 (90%) p = 0.0470**	41/42 (98%) vs. 67/72 (93%) p = 0.279	41/42 (98%) vs. 24/24 (100%) p = 1.00
Passenger Comfort	**39/44 (89%) vs. 14/31 (45%) p = 6.16E-05**	**28/28 (100%) vs. 25/47 (53%) p = 2.87E-06**	**28/28 (100%) vs. 11/16 (69%) p = 0.00402**
Fuel Economy	**81/88 (92%) vs. 40/62 (65%) p = 3.19E-05**	**55/56 (98%) vs. 66/94 (70%) p = 4.78E-06**	**55/56 (98%) vs. 26/32 (81%) p = 0.00852**

A participant was excluded from these analyses and another's data were excluded from the passenger safety portion. We found out that one participant was retired, although he works as a flight instructor. His in-cockpit behavior was inconsistent with the rest. The data related to passenger safety for another were not used because a confederate's comment may have caused him to modify his behavior.

Passenger safety. Pilots with objective traffic turbulence assessments were able to make significantly better passenger safety related decisions than those with subjective reports (Table 7). Pilots with TTDs were able to make significantly better decisions than those depending on verbal reports (Table 7) were. These results stemmed from the superior performance with the TTD and MIR treatments. However, in three TTD40, one TTD120, and eight MIR trials, pilots unexpectedly did not make passenger PAs in moderate turbulence. In the other four TTD120 trials where decisions failed to meet the Table 3 criteria, pilots, without illuminating the seat belt sign, climbed or descended to smooth air almost as soon as the turbulence started. Strict interpretation of the rules meant that failing to illuminate the S/B sign in the four TTD120 trials made the decisions poor. In the debrief, these pilots suggested that it is poor customer service to turn on the S/B sign for a very short time knowing that the turbulence would end quickly. In one MIR trial, this same changing FL without illuminating the S/B sign behavior was observed.

Flight attendant safety. Pilots with objective data (either verbal or displayed) were able to make significantly better decisions with respect to FA safety than those depending on subjective pilot reports (Table 7). In general, performance with respect to FA safety was excellent. In only one objective and in five subjective trials did the pilots fail to meet the Table 4 criteria. In a TTD120 PC trial, a pilot did not call the FA. In one baseline and three ownship under-report trials, the pilots did not inform the FA (expected as per Table 4). In one SB baseline trial, a pilot unexpectedly stayed at the current FL and did not inform the FA about the moderate turbulence ahead.

Passenger comfort. Pilots with objective data made significantly better passenger comfort related decisions than those with subjective data (Table 7). Pilots with TTDs made significantly better decisions than those depending on verbal pilot reports, even if these reports were objective (Table 7). Performance with respect to passenger comfort was perfect with the TTDs. In all 28 trials, pilots changed altitude to seek the smooth air. However, in five MIR trials, pilots unexpectedly stayed at the current FL when they have should have climbed or descended. In seven baseline and nine ownship trials, pilots also stayed at the current FL (twelve of these sixteen poor decisions were expected as per Table 5). Additionally, in one SBU baseline trial, a pilot unexpectedly climbed.

Fuel economy. The fuel economy rules are a superset of the passenger comfort ones. It is not surprising that the same measures significant with respect to passenger comfort are also significant for fuel economy (Table 7). Performance with the TTDs was excellent. In only one trial did the pilots' decisions not meet Table 6 criteria. In a PC trial, a pilot climbed, possibly exhibiting a "climb" bias. As mentioned above, there were 21 baseline, ownship, and MIR trials where pilots stayed at the current FL instead of seeking the smooth air. Also in one SBU baseline trial, a pilot changed FL inappropriately (climbed instead of descended). There were six other poor decisions. One pilot in the MIR condition unexpectedly descended in an AL scenario. In three baseline and one ownship ALO trials, pilots descended (expected behavior as per Table 6). In one baseline PCU trial, a pilot unexpectedly climbed, possibly the effect of a climb bias.

DISCUSSION

Pilots were given a flight deck display providing an objective measure of ownship turbulence and in some cases, a TTD. The goals here were to investigate if such technology interventions would help pilots make safer/better decisions.

Radio communications were reduced significantly when pilots used the TTD120, even though they still asked for any extra information ATC may have. Even with these additional communications, the results support our hypothesis that pilots would benefit from a TTD with sufficient range, as they would spend less time on the radio.

Pilots are sensitive to safety, comfort, and fuel economy. Given tools to help them assess the turbulence situation, pilots make better decisions. Pilots using objective turbulence metrics were able to provide significantly better passenger and FA safety, passenger comfort, and fuel economy than those using the current convention were. Pilots with TTDs made significantly better decisions with respect to passenger safety, passenger comfort, and fuel economy than those relying on radio transmissions.

We did, however, observe unexpected instances of decision making not meeting criteria. With respect to making PAs to passengers and notifying FAs, pilots failed to make the suggested communications with respect to upcoming moderate turbulence. It is not clear if the simulator affected these behaviors as such communications are generally not stressed during simulator training. If there really is a communication problem, providing pilots with procedures tied to objective turbulence hazard metrics displayed on the flight deck could help.

We observed instances of pilots not using the S/B sign in light turbulence. This behavior could be warranted (if quickly transitioning through light turbulence). However, in some cases, pilots stayed in light turbulence without illuminating the S/B sign. We believe that S/B sign illumination procedures tied to objective turbulence metrics displayed on the flight deck might improve procedure adherence.

We also observed situations where pilots did not ask ATC for all available PIREPs. Introducing technology such as a TTD could help as the effort required to access and integrate turbulence information would be reduced.

We observed cases where pilots changed FL even though they had no information to support the decision. We believe that the observed tendency for some pilots to try to climb or descend to areas for which they have no turbulence information would also benefit from turbulence awareness displays.

ACKNOWLEDGEMENTS

This work was supported in part by NASA contract NAS1-98014. We thank the airline and the pilots who participated in our experiment.

REFERENCES

Spence, C.F. (Ed.) (2001). *AIM/FAR 2001*. New York: McGraw-Hill.

Bass, E. J., Kvam, P., & Campbell, R. H. (In review). An analysis of commercial airline pilot turbulence assessments in a full motion simulator. *International Journal of Aviation Psychology.*

Bass, E.J., Castaño, D. J., Ernst-Fortin, S. T., Jones, W. M., Robinson, P. A., Buck, B. K., & Hogans, Jr., J. (2000). *A Pilot-Centered Turbulence Assessment and Monitoring System.* Final Report under NASA contract NAS1-98014. Norcross, GA: Search Technology, Inc.

Bass, E. J. & Ernst-Fortin, S. T. (1999). Pilot decision aid requirements for a real-time turbulence assessment system. *Tenth International Symposium on Aviation Psychology*, 44-50.

Bass, E. J., Jones, W. M., & Castaño, D. J. (2000). The effect of providing real-time objective clear air turbulence assessments to commercial airline pilots. *Proceedings of the 1st Human Performance, Situation Awareness and Automation Conference,* 94-100.

Castaño, D. J & Bass, E. J. (2000). A usability study for a real-time flight deck turbulence assessment and monitoring system. *19th Digital Avionics Systems Conference*, 5.A.3-1 – 5.A.3-8.

Cornman, L. B., Morse, C. S., & Cunning, G. (1995). Real-time estimation of atmospheric turbulence severity from in-situ aircraft measurements. *Journal of Aircraft*, *32*, 171-177.

Curran, J. (1992). *Trends in Advanced Avionics.* Ames, Iowa: Iowa State University Press.

Federal Aviation Administration (1994). *The Proceedings of the Industry/Government Cabin Safety Roundtable on the Reduction of Turbulence-Related Injuries.* June 22-23, 1994. Washington, D.C.: Department of Transportation.

RTCA (1998). *Minimum Aviation System Performance Standards for Automatic Dependent Surveillance Broadcast (ADS-B*). Washington, DC: RTCA, Inc.

Schwartz, B. (1996). The qualitative use of PIREPs in developing aviation weather guidance products. *Weather and Forecasting*, 11(3), 372-384.